基于转录组测序的枫香 EST-SSR 引物开发及有效性评价

李辉[1,2]，冯源恒[2]，唐生森[2]，杨章旗[1,2*]
（1. 广西师范大学 生命科学学院，广西 桂林 541006；2. 广西壮族自治区林业科学研究院 国家林业和草原局马尾松工程技术研究中心 广西马尾松工程技术研究中心 广西优良用材林资源培育重点实验室，南宁 530002）

摘　要：枫香是广西重要的乡土树种之一，有着较高的用材、观赏及药用价值，该文基于枫香转录组测序结果进行 EST-SSR 引物开发，可为枫香群体遗传学研究的开展提供有效、可靠的分子标记工具，对于该树种的保护、开发和利用有着重要的意义。研究基于转录组测序技术，检测枫香 SSR 位点并设计引物，通过 PCR 扩增和聚丙烯酰胺凝胶电泳筛选出具有较高多态性的枫香 EST-SSR 引物，并通过对 1 个枫香天然群体进行遗传多样性分析验证这些引物实际应用效果。结果表明：（1）基于转录组测序得到的枫香 SSR Unigene 中共发掘到 23 777 个 SSR 位点，单核苷酸重复类型 SSR 位点占总位点比例最高（46.54%）。重复次数上，5~12 次之间的 SSR 位点占比最高（72.36%）。（2）共开发出 262 对 SSR 引物，有效扩增率为 53.1%，最终筛选出扩增稳定、条带清晰的引物 18 对。（3）多态性检测结果表明，所有位点均具有多态性；天然群体遗传多样性结果显示，该天然群体中等位基因数量（*Na*）、有效等位基因数量（*Ne*）、Shonnon 信息指数（*I*）、观察杂合度（*Ho*）变化范围分别为 2~4、1.112 8~2.609 6、0.208 9~1.112 7 和 0.275 9~1.000 0，平均值分别为 2.333 3、1.957 4、0.708 5 和 0.722 6。综上所述，在枫香中占优势的 SSR 位点重复类型和重复基序与其他物种基本相同；本研究所开发的 18 对枫香 EST-SSR 引物可以满足开展枫香群体遗传学研究的需要，为后续枫香种群遗传多样性研究提供了丰富的标记引物。

关键词：枫香，转录组，SSR 分子标记，引物开发，遗传多样性
中图分类号：Q943　　　　**文献标识码：**A

Development and validity evaluation of *Liquidambar formosana* EST-SSR primers based on transcriptome sequencing

LI Hui[1,2], FENG Yuanheng[2], TANG Shengsen[2], YANG Zhangqi[1,2*]
(1. *College of Life Sciences*, *Guangxi Normal University*, Guilin 541006, Guangxi, China; 2. *Guangxi Forestry Research Institute*, *Engineering Research Center of Masson Pine of State Forestry Administration*, *Engineering Research Center of Masson Pine of Guangxi*, *Guangxi Key Laboratory of Superior Timber Trees Resource Cultivation*, Nanning 530002, China)

Abstract: *Liquidambar formosana* is one of the important native tree species in Guangxi, which has high timber, ornamental and medicinal value. This study designs and develops EST-SSR markers of *L. formosana* based on data from the transcriptome sequencing. The purpose is to provide an effective and reliable molecular marker tool for studies of population genetics of this species, it is of great significance to the protection and utilization of all kinds of material resources of the tree. Simple sequence repeats were selected from the transcriptome data in *L. formosana*, and then primers were developed, and screened out by PCR amplification and polyacrylamide gel electrophoresis for high polymorphism, and the efficiency was tested by using 30 individuals from of a wild *L. formosana* population. The results were as follows: (1) A total 23 777 SSRs were obtained by searching SSR unigenes from transcriptome data. The repeat type of the SSR loci in *L. formosana* was dominated by mononucleotide repeats (46.54%). The highest proportion of SSR loci (72.36%) was between 5 and 12 times. (2) A total of 262 pairs of SSR primers were developed, and among them, 139 pairs amplified the target bands with a

基金项目：广西科技基地和人才专项（桂科 AD19254004）；八桂学者专项（2019A26）[Supported by Guangxi Science and Technology Project (GuikeAD19254004); Bagui Scholar Program of Guangxi (2019A26)]。
第一作者：李辉(1995 -)，硕士研究生，研究方向为生物化学与分子生物学，（E-mail）lihui2332022@163.com。
***通信作者：**杨章旗（1964 - ），博士，教授级高级工程师，研究方向为林木遗传育种，（E-mail）yangzhangqi@163.com。

success rate of 53.1%, and 18 pairs of them that can be used to steadily obtain clear bands were finally identified. (3) The polymorphism detection showed that all sites had a high degree of polymorphism. The number of alleles, effective alleles, Shannon diversity index, observed heterozygosity and fixed index of the *L. formosana* populations ranged in 2−4, 1.112 8−2.609 6, 0.208 9−1.112 7 and 0.275 9−1.000 0, the average values were 2.333 3, 1.957 4, 0.708 5 and 0.722 6, respectively. In conclusion, the dominant SSR repeat type and repeat motif in *L. formosama* are basically the same as those in other species, the developed EST-SSR markers can meet the needs of population genetic studies of *L. formosana*, which provides abundant primers for the study of genetic diversity of *L. formosana.*

Key words: *Liquidambar formosana*, transcriptome, SSR molecular markers, primers development, genetic diversity

枫香（*Liquidambar formosana*）是金缕梅科（Hamamelidaceae）枫香属（*Liquidambar*）的彩叶落叶乔木，主要分布于秦岭淮河一线以南，广西全境都有分布。叶片春夏呈青绿色，秋季呈深红色，具有较高的观赏价值。作为我国传统药用植物之一，枫香树皮可用于通经活络，果子对于消化不良等病症有着明显的作用，叶片有着清热解毒的功效，更是广西壮族“三月三”传统节日中制作五色米饭不可或缺的重要材料。枫香还有着较高的用材价值，可用于制作家具或建材。除此之外还有着维护地力的重要价值（范绍斌等，2021），其适应力强、耐贫瘠，对于一些有毒气体如 Cl_2 、SO_2 有着较强的抗性（赖玖鑫，2020），常作为生态结构、林种结构调整的重要树种，因此有着较高的研究价值。

近些年来，转录组测序（RNA-seq）技术被广泛应用到分子生物学领域，逐渐成为分子生物学研究中最常用的技术(易浪波等，2017)，转录组测序技术能够快速有效地获得基因序列，并从整体上揭示特定阶段物种基因功能及结构（刘佳奇等，2020），通过对转录组测序数据进行挖掘分析，可以得到 SSR 位点分布特征，从而为后期 SSR 引物的筛选和设计提供数据参考(叶鹏等，2019)。简单重复序列（simple sequence repeats，SSR)又称为微卫星 DNA，指由 1~6 个核苷酸为重复单元串联组成的长达几十个核苷酸的重复序列，SSR 在真核生物整个基因组内都有存在，有着数量丰富、多态性高、重复性好、特异性强的优点，被广泛应用于物种遗传多样性分析、遗传连锁图谱的构建（张利达和唐克轩，2010）、基因定位和分子标记辅助育种等研究(杨雄等，2021)，被视为检测群体遗传多样性和群体间遗传分化最有效标记之一。按其来源可分为基因组 SSR (G-SSR) 和基于表达序列标签（expressed sequence tag，EST）的 SSR 两类。其中，G-SSR 标记基于基因组序列，其开发过程繁复、成本高、效率低；EST-SSR 标记基于表达序列标签，除具有 G-SSR 标记的优点外，还具有序列保守性较高，在植物同属不同物种间有较好通用性的优点(王丹丹和杨东霞，2017)。

枫香作为我国重要乡土树种之一，国内主要开展了生理性质分析（尹国平等，2021）、人工林培育（韦理电等，2016）及新品种（王建军等，2015）认定等工作。关于枫香群体遗传学方面研究相关报道较少，其原因主要是缺乏足够的分子标记，以往对枫香遗传多样性的研究主要采用显性分子标记，如黄敏等（2021）探索了最适的枫香叶片 DNA 的提取方法，最优 ISSR 引物和最佳 ISSR-PCR 扩增反应体系；毕则鑫等（2010）通过 ISSR 分子标记技术分析了枫香自然群体遗传多样性；李芳芳（2015）优化了 DNA 提取及 SRAP-PCR 反应体系的方法等。相较而言，ISSR 标记实验重复性较 SSR 标记差，在多等位位点的检测、杂合子的区分及基因分型等方面无法具备 SSR 标记的功能；而 SRAP 标记的扩增谱带过多，当用多个引物组合来构建品种指纹图谱时，指纹的识别和鉴定相较 SSR 标记更为困难（刘振等，2014），SSR 分子标记可揭示的遗传信息相对较高，明显优于其他传统分子标记(刘粉粉等，2021），但当前关于枫香 SSR 引物开发的报道相对较少。前人的研究中主要通过磁珠富集法（顾辛韵，2016）和从 NCBI 获取数据的方法（孙荣喜，2017）开发枫香 SSR 引物，该方法都存在着获得位点信息较少，开发引物难度大，开发数量较少的问题，由此使得当前枫香 SSR 引物数量还不足以支撑更深入地进行枫香群体遗传学研究。

本研究以河池市 7 个县及贵州荔波县为研究区域，依托转录组测序技术获得丰富且全面的枫香 SSR

位点信息，进行枫香 EST-SSR 引物的筛选和设计，随后采用 PCR 扩增及聚丙烯酰胺凝胶电泳的方法筛选出多态性好，扩增稳定的枫香 EST-SSR 引物，并通过对野生群体的遗传多样性分析以验证引物的应用效果，拟讨论以下问题：（1）枫香 SSR 位点的具体分布及主要的序列特征；（2）枫香 SSR 引物的开发效率如何，影响开发效率的原因有哪些；（3）开发出的枫香 SSR 引物多态性情况及在枫香遗传多样性分析当中的应用效果。以期为今后枫香分子标记辅助育种、功能基因标记等研究提供基础依据，为枫香群体遗传学及遗传育种研究提供可靠研究工具，为进一步开展广西境内枫香种质资源遗传多样性评价、枫香分子植物辅助育种等研究提供参考，为枫香育种材料遗传多样性及亲缘关系分析奠定基础。

1 材料与方法

1.1 研究材料

从广西河池市环江县随机选择一个枫香天然群体，采集其新鲜的红、黄、绿三种颜色的叶片分别装入 50 mL 离心管，并放入液氮中冷冻封存，交由北京诺禾致源生物信息科技有限公司进行转录组测序。经查阅文献并实地考察后，从广西河池市 7 个县及贵州荔波县各选择 1 个枫香天然群体随机采集 1 个枫香样本（A1~A8）用作引物的初步筛选，另在 8 个天然群体中各采集 2~4 个枫香样本（各采样单株之间相距＞50 m），共计 24 个枫香样本（B1~B24）用作引物的复筛及多态性检测材料，从而保证实验材料能够覆盖该试验区更多天然群体及个体；此外，在河池市环江县另选择一个枫香天然群体采集 30 个样本（C1~C30）用作遗传多样性分析的实验材料。材料具体来源如表 1 所示。

表 1 枫香实验材料来源地

Table 1 Sources of *Liquidambar formosana* experimental materials

编号 No.	材料来源 Source of materials	经纬度 Longitude and latitude
A1，B1~B4	广西河池南丹县 Hechi Nandan County, Guangxi	107°20′56″ E 25°23′38″ N
A2，B5~B6	广西河池天峨县 Hechi Tian’e County, Guangxi	107°11′30″ E 25°08′06″ N
A3，B7~B8	广西河池凤山县 Hechi Fengshan County, Guangxi	106°54′44″ E 24°31′56″ N
A4，B9~B10	广西河池东兰县 Hechi Donglan County, Guangxi	107°29′10″ E 24°27′38″ N
A5，B11~B14	广西河池金城江区 Hechi Jinchengjiang District, Guangxi	108°33′36″ E 24°34′58″ N
A6，B15~B18	广西河池环江县 Hechi Huanjiang County, Guangxi	108°33′31″ E 25°17′16″ N
A7，B19~B21	广西河池宜州区 Hechi Yizhou District, Guangxi	108°45′52″E 24°34′17″N
A8，B22~B24	贵州荔波县 Libo County, Guangxi	108°07′55″E 25°15′25″N
C1~C30	广西河池环江县 Hechi Huanjiang County, Guangxi	108°33′02″ E 25°18′41″ N

1.2 枫香转录组测序

主要从以下方面对测序过程进行把控。（1）对样品质量检测：琼脂糖凝胶电泳分析 RNA 降解程度以及是否有污染；Nanodrop 检测 RNA 的纯度（OD260/280 比值）；Qubit 对 RNA 浓度进行精确定量；Agilent 2100 精确检测 RNA 的完整性，从而保证 RNA 的完整性和总量。（2）使用 Oligo（dT）磁珠富集

的方法构建全长转录组文库，在 PCR 扩增富集合成的 cDNA 时，通过循环优化确定 PCR 的最佳条件，从而保证文库构建质量，经检验合格后，根据文库的有效浓度及数据产出需求运用 PacBio Sequel 平台进行测序。从而保证最终得到的 SSR 位点信息有效可靠。

1.3 枫香 EST-SSR 标记开发

对转录组测序挖掘到的枫香转录组中 SSR 位点进行筛选，筛选的标准是序列长度单位范围为 16~28 bp，2 个碱基单元的重复次数大于 9 次，3~6 个碱基单元的重复次数为 5 次或大于 5 次，单碱基单元排除在外，上、下游引物退火温度差不超过 3 ℃等条件开发设计 SSR 引物，交由广州艾基生物有限公司进行引物合成。

1.4 枫香基因组 DNA 提取

本研究使用快捷型基因组 DNA 提取系统提取枫香 DNA（王茜等，2013），随后用分光光度计对提取后的 DNA 进行纯度与浓度检测，OD260/OD280 值应为 1.7~1.9。随后通过 1%琼脂糖电泳检验 DNA 质量，将符合要求的模板 DNA 稀释到 50 ng·μL^{-1}，放置于-20 ℃冰箱中备用。

1.5 PCR 扩增

PCR 扩增反应总体积为 10 μL，其中包含 2 μL DNA（50 ng·μL^{-1}），1 μL PCR Buffer（10×），0.2 μL dNTPs，上下游引物各 0.25 μL，0.07 μL Taq DNA Polymerase（5 U·μL^{-1}），及 6.23 μL 的 ddH_2O，该扩增反应在 PCR 仪上完成。SSR-PCR 扩增的程序：94 ℃预变 4 min，94 ℃变性 15 s，55 ℃复变性 15 s，72 ℃延伸 30 s，30 个循环，72 ℃延伸 20 min，12 ℃保存。

1.6 引物有效性及多态性检测

将 PCR 扩增后的产物通过 8%聚丙烯酰胺凝胶电泳进行分离，实验步骤为：使用移液枪吸取 1.2 μL 的滴加了溴酚蓝指示剂的扩增引物点在聚丙烯酰胺胶板上，凝胶首尾点 1 μL 50~500 bp Marker，在 240 V 恒压下电泳，通过溴酚蓝指示剂调整电泳时间。电泳结束后将凝胶使用 ddH_2O 清洗 2 次，放入固定液中震荡 10 min，随后取出凝胶放入超纯水中震荡清洗 2~3 次，每次 2 min。清洗后将胶板放入银染液中进行银染，震荡 7 min；银染后再次放入蒸馏水中震荡清洗 2~3 次，每次 2 min。将清洗后的凝胶放入显影液中震荡，直至胶板上出现清晰条带为止。用 ddH_2O 清洗凝胶 2 次后读条带并拍照保存。

1.7 数据统计

人工进行条带判读，通过是否扩增出有效条带判断引物有效性。对具有有效性的引物，其不同扩增产物按条带长度从大到小用 A, B, C...... 进行编号，汇总后通过 Popgene 32 软件计算以下参数：（1）观测等位基因数目(observed number of alleles，*Na*)；（2）有效等位基因数目 (effective number of alleles，*Ne*)（Hartl et al., 1989）；（3）Shannon 多样性指数（Shannon’s information index，*I*）（Shannon et al., 1949）；（4）观测杂合度（observed heterozygosity，*Ho*)；（5）期望杂合度(expected heterozygosity，*He*）（Nei et al., 1973）。通过 PIC-CALC 软件计算 *PIC*（polymorphism information content）值（Botstein et al., 1980），*PIC* 值指多态信息含量，是判断 SSR 引物多态性的重要指标，当 $PIC \leqslant 0.25$ 时，该位点为低度多态性位点；当 $0.25<PIC<0.5$ 时，该位点为中度多态性位点；当 $PIC \geqslant 0.5$ 时，该位点为高度多态性位点。计算 *PIC* 值所用到的公式（杨国忠等，2004）如下：

$$PIC = 1 - \sum_{i=1}^{n} P_i^2 - \sum_{i=1}^{n-1} \sum 2P_i^2 P_j^2 \quad (1)$$

式中：*PIC* 指多态信息含量；P_i 和 P_j 分别表示第 *i* 个和第 *j* 个等位基因频率；*n* 表示等位基因数。

引物的有效扩增率指能够有效扩增的引物数量与引物总数量的比值。引物的多态性比率指具有能够有效扩增且多态性的引物数量与引物总数量的比值。

2 结果与分析

2.1 枫香转录组中 SSR 位点的分布特点

通过枫香转录组测序共挖掘到 23 777 个枫香 SSR 位点，平均每 3.13 kb 存在一个 SSR 位点。通过统

计发现，枫香转录组中 SSR 主要有单核苷酸及二到六个核苷酸等六个简单重复单元类型，除此之外还有少量复杂重复类型（表 2），其中单核苷酸类型占 SSR 位点总数的 46.54%，其次则是二核苷酸重复类型，占 SSR 位点总数的 33.10%，三、四、五、六核苷酸类型分别占 SSR 位点总数的 17.80%、1.17%、0.49%、0.88%。单核苷酸类型中出现频率最高的是 T/A（66.81%），其次是 A/T（24.99%）；在二核苷酸类型当中，出现频率最高的是 CT/AG（24.7%），较之稍少的是 TC/GA（23.72%）；在三核苷酸类型当中，出现最多的则是 CAG/CTG（6.17%），次之的是 GAA/TTC（5.08%）；在四核苷酸类型中出现最多的是 TTTA/TAAA（17.2%）；在五核苷酸类型中，CTTTT/AAAAG（8.62%）出现的频率最高；在六核苷酸类型中出现最多的则是 ACCAGC/GCTGGT（4.72%）。综上所述，枫香 SSR 位点以单核苷酸、二核苷酸为主，共占 SSR 位点总数的 79.64%，适用于设计引物的三到六核苷酸重复类型位点总数仅占位点总数的 19.46%，限制了在此基础上开发枫香 SSR 引物的数量。

表 2 枫香转录组中 SSR 重复单元的分布特征

Table 2 Distribution of the SSR repeat motifs in *Liquidambar formosana* transcriptome

核苷酸重复类型 Type of nucleotide repeat	SSR 数量 No. of SSR	比例 Percentage (%)	主要重复基元（数目） Main repeat motif (No.)
单核苷酸 Mononucleotide	11 067	46. 54	T/ A （7 394） A/ T （2 766）
二核苷酸 Dinucleotide	7 871	33. 10	CT/ AG（1 944） TC/ GA （1 867） AG/ CT （1 162）
三核苷酸 Trinucleotide	4 232	17. 80	CAG/ CTG（261） GAA/ TTC（215） TCT/ AGA （176） GCA/ TGC（175）
四核苷酸 Tetranucleotide	279	1. 17	TTTA/ TAAA （48） ATTT/ AAAT（27）
五核苷酸 Pentanucleotide	116	0. 49	CTTTT/ AAAAG（10） TTCTG/ CAGAA （9） TCTCT/ AGAGA （9）
六核苷酸 Hexanucleotide	212	0. 88	ACCAGC/ GCTGGT（10） ATGGGT/ ACCCAT （9） TCGCTC/ GAGCGA（9）
总数 Total	23 777		

SSR 重复次数（表 3）在 5~78 次之间，主要集中在 5~24 次之间，其中重复次数数量在 SSR 总位点数量中占比最多的是 5~8 次，共有 8 787 个，占 SSR 位点总数的 36.96%，排在第二位的则是 9~12 次重复，共有 8 417 个，占 SSR 位点总数的 35.40%，重复次数在 13~16 次、17~20 次和 21~24 次的 SSR 位点数量分别为 3 934 个（16.55%）、1 830 个（7.70%）和 598 个（2.52%），重复次数大于 24 次的最少，共有 211 个，仅占 SSR 位点总数的 0.89%。在所有的重复次数当中，10 次重复所占比例最大，有 3 570 个，占 SSR 位点总数的 15.01%。稍次之的是 6 次重复，共有 3 220 个，占 SSR 位点总数的 13.54%。SSR 位点的多态性由 DNA 复制过程中的滑动错配产生，SSR 的重复次数越多，该位点的多态性往往越高，枫香 SSR 重复次数以 5~8 次为主，基本满足设计开发 SSR 引物的要求。

表 3 SSR 重复次数

Table 3 SSR repetitions

重复次数 Repetitions	5~8 次 5−8 times	9~12 次 9−12 times	13~16 次 13−16 times	17~20 次 17−20 times	21~24 次 21−24 times	＞ 24 次 ＞ 24 times
数量 Number	8 787	8 417	3 934	1 830	598	210
比例 Percentage (%)	36. 96	34. 40	16. 55	7. 70	1. 52	0. 89

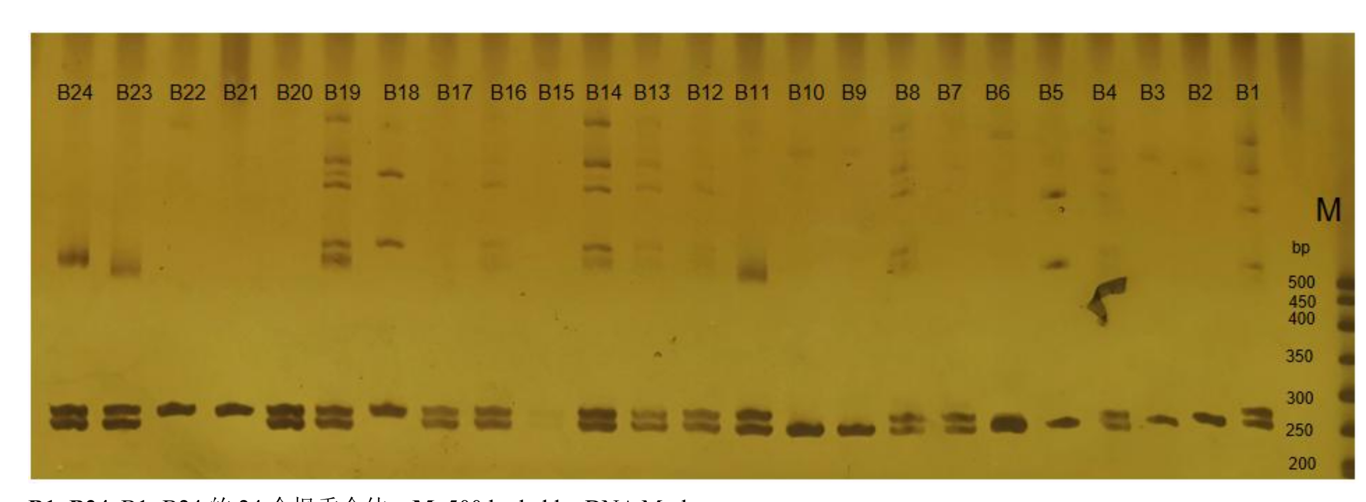
2.2 枫香 EST-SSR 引物有效性筛选

如表 4 所示，用 8 个枫香样本（A1~A8）DNA 对 262 对引物进行简捷有效地初次筛选，有 139 对引物能够有效扩增，有效扩增率为 53.1%，且所设计的不同核苷酸重复类型的引物有效扩增率基本相同。因此可认为本实验所设计合成的各核苷酸重复类型引物在枫香群体具有一定的适用性（杨雄等，2021）。

表 4 枫香 EST-SSR 引物扩增结果

Table 4 Amplification results of EST-SSR primers of *Liquidambar formosana*

SSR 核心类型 Type of SSR core	引物对数目 No. of primer pairs	有效引物数目 Number of effective primers
二核苷酸 Dinucleotide	10	5（50.0%）
三核苷酸 Trinucleotide	146	78（53.4%）
四核苷酸 Tetranucleotide	76	40（53.6%）
五核苷酸 Pentanucleotide	30	16（53.3%）
总计 Total	262	139（53.1%）



B1~B24. B1~B24 的 24 个枫香个体；**M**. 500 bp ladder DNA Marker。

B1-B24. DNA of 24 *Liquidambar formosana* individuals from B1 to B24; **M**. 500 bp ladder DNA Marker.

图 1 24 份枫香材料扩增结果（LF229）

Fig. 1 Amplification of 24 *Liquidambar formosana* using primer LF229

2.3 枫香 EST-SSR 引物多态性分析结果

筛选出的 139 对有效性引物中，能够扩增出多态性条带的引物有 41 对，多态性比率为 15.64%，说明

枫香 SSR 序列的保守性相对较强。选取其中 18 对扩增稳定，条带清晰，有多态性的引物（表 5），对抽取的 24 个样本（B1~B24）DNA 进行多态性检测，并用 PopGene32 软件进行分析。

多态性分析结果显示，18 个位点的观测等位基因数量（*Na*）介于 2~5 之间，平均值为 3；*PIC* 值介于 0.30~0.72 之间，其中 10 个位点的 *PIC*＞0.50，具有高度多态性，其余 8 个位点的 *PIC*＞0.25，具有中度多态性，不存在低多态性位点。经筛选出的 18 对枫香 SSR 引物具有较高的多态性。

表 5 18 对 SSR 引物多态性信息

Table 5 Information on polymorphism of 18 pairs of SSR primers

引物编号 Primer No.		引物序列（5'-3'） Primer sequence （5'-3'）	重复基元 Repeat motif	退火温度 Annealing temperature （℃）
LF53	F:	AGGAACCCAATTTGACCCCT	(TGG) 6	60.0
	R:	CCACCCACCAATGCTCAGAT		
LF76	F:	CAGCACTTCCATAAAGCCGC	(CCAT) 5	60.0
	R:	GTTCTGACGCCTTCTGGTGA		
LF89	F:	TGAAATTTTGCGTGGGCTGG	(GGCT) 5	57.3
	R:	ACAAGTGGGCTCAAATGAGT		
LF108	F:	GCAATAGCCGTTCCCTTTGC	(TTCT) 5	60.0
	R:	CACGGTCTCACTGCCTGATT		
LF111	F:	CGGAGCTCGTGATTTCAGGT	(TTTC) 5	60.6
	R:	AGAACCGGTTCCTCGTAACA		
LF112	F:	ATCAGTGTGTAGGCCGAACC	(TTTG)5	58.6
	R:	AGCCAATCACCACACACCAA		
LF128	F:	CACACCACTCCACTCCGATC	(CCA) 7	60.0
	R:	ATGTTGAGGGAGCGAAGGTG		
LF129	F:	GTATGATCGCCCTCCCACTG	(CCG) 7	60.0
	R:	ACAGAAGCTGAGGTTGAGGC		
LF171	F:	ACTCACACTCCCACACTCCA	（CCT) 8	60.0
	R:	AAACTCGAAGACCCAGCGAG		
LF177	F:	CCACTGCCTCAGGGTTTCTT	(GAAG) 6	60.0
	R:	CTTCACCTGGAAGTGGGTCG		
LF188	F:	TTCTGTGACTGGGCAATGGG	(TAA) 8	59.8
	R:	CGAGTCCTCACATTCCACGT		
LF190	F:	TGGTGCCTTCTTGGACCTTT	(TATT) 6	60.0
	R:	CACAACAAACGAGGAAGCCG		
LF206	F:	CAAATGCGAATGAGCGGGAG	(AAAGG) 5	60.0
	R:	GGGCAAAAGGGACAGGATGT		
LF224	F:	ACAGGCTGTGGGATTAGCAC	(GAAGG) 5	60.0
	R:	GTCGTCAGAGTGGCCAGAAA		
LF229	F:	GCTACCGTTGCCTCCTTCAT	(TTTCC) 5	59.8
	R:	TGGTCTTGGAGCAGATGACG		
LF242	F:	TCTCATCGCACTCCTAGGGT	(GAA) 9	60.0
	R:	GTGAGTGGAGAGGGGGAGAT		
LF245	F:	GTGCTGGATCTCTGGCACAA	(GCA) 9	59.9
	R:	GACCAATGTATCCCGGGGAC		

LF247	F： R：	GGAAGTACCACTCCTTCGGC GAAACTCCGAATTCGACGCG	(GCG) 9	60.0

2.4 枫香天然群体 EST-SSR 遗传分析有效性评价

用这 18 对引物对一个枫香天然群体的 30 个样本（C1~C30）DNA 进行遗传多样性分析，扩增结果显示（表 6），18 个位点共扩增出 42 个等位基因数，观测等位基因数（*Na*）介于 2~ 4 之间，平均值为 2.333 3；有效等位基因数（*Ne*）介于 1.112 8~2.609 6 之间，平均值为 1.957 4；Shannon 指数（*I*）介于 0.208 9~1.112 7 之间，平均值为 0.708 5；观测杂合度（*Ho*）介于 0.275 9~1.000 0 之间，平均值为 0.722 6；期望杂合度（*He*）介于 0.370 6~0.896 8 之间，平均值为 0.522 8。该结果表明，开发的 18 对 SSR 引物可用于枫香群体遗传多样性评价。

表 6 18 个位点的遗传多样性

Table 6 Genetic diversity of the 18 loci

位点 Locus	等位基因数目 Number of allele (*Na*)	有效等位基因数 Number of effective allele (*Ne*)	Shannon 指数 Shannon's diversity index (*I*)	Nei's 基因多样性指数 Nei's gene diversity （*H*）	观测杂合度 Observed heterozygosity (*Ho*)	期望杂合度 Expected heterozygosity (*He*)
LF53	2	1.997 6	0.692 6	0.499 4	0.862 1	0.491 8
LF76	3	2.166 4	0.914 1	0.538 4	0.880 0	0.450 6
LF89	4	2.609 6	1.112 7	0.616 8	0.800 0	0.370 6
LF108	2	1.773 8	0.627 9	0.436 2	0.714 3	0.555 8
LF111	3	2.179 4	0.894 3	0.541 2	0.555 6	0.448 6
LF112	3	2.453 7	0.967 4	0.592 5	0.730 8	0.395 9
LF128	2	1.989 8	0.690 6	0.497 4	0.928 6	0.493 5
LF129	2	1.112 8	0.208 9	0.101 4	0.964 3	0.896 8
LF171	3	2.179 4	0.878 9	0.541 2	0.592 6	0.448 6
LF177	2	1.996 8	0.692 3	0.499 2	0.680 0	0.490 6
LF188	2	1.990 5	0.690 8	0.497 6	0.275 9	0.493 6
LF190	2	1.997 5	0.692 5	0.499 4	0.392 9	0.491 6
LF206	2	1.997 3	0.692 5	0.499 3	0.925 9	0.491 3
LF224	2	1.337 6	0.419 5	0.252 4	0.777 8	0.742 8
LF229	2	1.997 3	0.692 5	0.499 3	0.407 4	0.491 3
LF242	2	1.975 6	0.687 0	0.493 8	1.000 0	0.496 9
LF245	2	1.997 3	0.692 5	0.499 3	0.777 8	0.491 3
LF247	2	1.480 2	0.505 5	0.324 4	0.740 7	0.669 5
平均值 Mean	2.333 3	1.957 4	0.708 5	0.468 3	0.722 6	0.522 8

3 讨论与结论

3.1 枫香转录组 SSR 位点分布特征

SSR 标记具有在植物种间通用性的优点，同属植物中移植已有标记进行开发 SSR 的研究已不少见(杨世鹏等, 2018)，但目前可用的枫香 SSR 引物较少，仅有报道的顾辛韵（2016）采用磁珠富集法开发的 10

对引物及孙荣喜（2017）通过 NCBI 数据库信息开发了 11 对引物。磁珠富集法主要是通过构建微卫星富集文库，经阳性克隆、筛选、测序最终得到 SSR 位点信息，该方法虽然在一定程度上提高了 SSR 引物开发效率，但操作过程繁琐且需要构建和筛选基因组文库，开发难度较大，获得的 SSR 位点信息有限；通过 NCBI 数据库开发 SSR 引物的方法使得研究者可以从中获得有效信息，节省了时间和试验成本，但当前数据库中的枫香序列数据有限，能够获得的 SSR 位点信息较少，且无法对枫香 SSR 位点信息分布特征进行详细地描述。随着高通量测序技术的发展，RNA-seq 测序的成本正逐渐降低，基于转录组数据开发 EST-SSR 分子标记，目前已经成为基因发掘及分子标记开发的重要技术手段。本研究首次通过转录组测序技术获得了相对较多且较全面的枫香 SSR 位点信息。本研究发现，作为第三纪的孑遗植物，枫香转录组中 SSR 基元重复类型较为丰富，包含了六种核苷酸重复的所有类型，高比例的单核苷酸重复序列（46.54%）是枫香 EST-SSR 的重要分布特征，其重复基元分布类型与大花四照花(*Cornus florida*) (刘佳奇等, 2020)和黄枝油杉（*Keteleeria calcarea*）相似（石远婷等，2021）。整体上看，枫香转录组中 SSR 位点的数量和种类较为丰富，为枫香 SSR 引物开发及后续研究提供了重要基础。

3.2 枫香 EST-SSR 位点开发效率

本研究所设计的 262 对枫香 EST-SSR 引物有效扩增率为 53.1%，祁雅等（2004）以及魏利斌等（2008）研究指出的，所设计 EST-SSR 引物的有效扩增率应在 60%～90%之间。本研究有效扩增率较低的原因可能有以下 3 个方面：（1）三至六碱基重复类型的 SSR 位点占比较低（总计 19.46%），使得可供筛选的 SSR 位点总量偏少；（2）部分引物设计不合理，实验中所用引物均根据二代测序结果所得，二代测序在拼接过程中可能会存在一些错误，因而导致有效扩增率相对较低；（3）本研究中开发的枫香 SSR 引物多态性比率较低（15.64%），主要是由于 SSR 的潜在多态性会随着 SSR 基元重复次数的增多而增加（刘佳奇等，2020），而枫香主要基元重复次数为 5~12 次，且主要基元以短重复类型为主，该 SSR 位点特征限制了枫香 SSR 的潜在多态性。内在原因主要是转录组 SSR 位点来自于序列相对保守的基因编码区，其多态性位点一般位于功能序列的微卫星重复片段上，从而导致转录组 SSR 多态性较低(杨世鹏等，2018)。

3.3 枫香 EST-SSR 引物多态性及群体遗传分析有效性评价

衡量 SSR 引物可用性及多态性丰富度可以采用多态信息含量（*PIC*）、观测等位基因数（*Na*）及有效等位基因数（*Ne*）等参数。在 *PIC* 指数方面，孙荣喜（2017）开发的 14 对枫香 SSR 引物中包括 6 个高度多态性位点，4 个中度多态性位点，4 个低度多态性位点。与前两者的研究相比，本实验中开发的 18 对枫香 SSR 引物中，10 对为高度多态性位点，8 对为中度多态性位点，多态信息含量（*PIC*）较高。在等位基因数（*Na*）方面，孙荣喜（2017）在全国范围内开展的中国枫香树遗传多样性及谱系地理研究中开发的 14 对枫香 SSR 引物，其观测等位基因数（*Na*）平均值为 3.258 0，有效等位基因数（*Ne*）平均值为 1.848 0；顾辛韵（2016）在千岛湖地区开展的片断化生境对枫香遗传多样性影响的研究中开发的 10 对 SSR 引物，观测等位基因数（*Na*）均值为 4.6，有效等位基因数（*Ne*）平均值为 2.130 0。本研究所开发 18 对 SSR 引物在环江群体中检测到的观测等位基因数（*Na*）平均值为 2.333 3，低于前两者，其原因是观测等位基因数（*Na*）在很大程度上依赖于抽取的样本大小，本研究采样范围及数量较前两者偏小；有效等位基因数（*Ne*）平均值（1.957 4），与前两者研究相近，而有效等位基因数（*Ne*）比杂合体频率和纯合体频率更好地度量群体遗传变异地增加，说明该环江群体遗传多样性较为丰富，上述讨论说明本研究所开发的 SSR 引物同样能够很好地反映出枫香群体遗传变异情况，可适用于枫香遗传多样性研究。

引物开发的目的是为了得到多态信息含量（*PIC*）高的 SSR 分子标记。前人研究中采用常规 SSR 分子标记开发技术获得的枫香序列有限，因此开发的枫香 SSR 引物数量较少、多态信息含量（*PIC*）不高，本研究首次通过转录组测序技术对枫香 SSR 位点信息进行了较为全面的发掘，可高效地获得枫香某一发育时期细胞全部转录本信息，并克服了步骤繁琐、时间和人力成本高的问题。由此自主设计枫香 EST-SSR 引物，并对其有效性及多态性进行验证，得到了 18 对尚未开发的多态信息含量（*PIC*）较高的枫香 EST-SSR 引物，有效性评价结果显示这 18 对枫香 SSR 引物均可用于后续枫香遗传多样性分析、枫香种质资源的保护研究中，在前人的基础上拓展了当前枫香 SSR 引物数量和种类。由群体验证结果可知，不同的

EST-SSR 分子标记在群体表达上存在一定的差异，说明不同位点所在基因的进化速率存在着差异，多态性较低的 SSR 位点多位于保守基因上（刘佳奇等，2020）。通过转录组测序技术开发的枫香 EST-SSR 引物除用于遗传多样性分析外，一方面可通过其位点特征回溯所在基因的固定或进化情况，以探究群体遗传分化问题；另一方面结合转录组信息可开展位点所在序列基因功能性研究，既可用于构建优良表型关联的特异 SSR 标记，还可为后续深入开展枫香叶色变化分子机理、枫香药用活性成分的合成等研究提供一定的帮助。对于枫香种质资源的开发、保存和利用有着重要的指导意义。

综上所述，本研究基于转录组测序技术获得了较为全面的枫香SSR位点信息，成功开发出18对枫香EST-SSR引物，对枫香中高度多态性SSR引物的数量提供了必要的补充，为解决当前枫香中高多态性SSR引物数量和种类不足以支撑更深入地进行枫香遗传多样性研究的问题提供了帮助，为枫香群体遗传学、分子辅助育种等研究提供了更多、更好的研究工具。

参考文献：

BI ZX, JIN ZX, LI JM, et al., 2010. Genetic Diversity in the natural populations of *Liquidambar fornosana* revealed by ISSR molecular markers [J]. Bull Bot Res, 30(1): 120-125. [毕则鑫，金则新，李钧敏，等，2010. 枫香自然种群遗传多样性的ISSR分析 [J]. 植物研究，30(1): 120-125.]

Botstein D, White R L, Skolnick M, et al., 1980. Construction of a genetic linkage map inman using restriction fragment length polymorphisms [J]. Am J Hum Genet, 32(3): 314-331.]

FAN SB, SU JY, FANG T, et al., 2021. Pathogen of *Liquidambar formosana* leaf spot. mycosystema [J]. https://kns.cnki.net/kcms/detail/11.5180.Q.20211228.1418.002.html. [范绍斌，苏济钰，方甜，等，2021. 枫香叶斑病病原菌 [J]. 菌物学报，https://kns.cnki.net/kcms/detail/11.5180.Q.20211228.1418.002.html.]

GU XY, 2016. Impacts of habitat fragmentation on genetic diversity and differentiation of *Liquidambar formosana* populations [D]. Shanghai: East China Normal University: 6-7. [顾辛韵，2016. 生境片断化对枫香（*Liquidambar formosana*）种群遗传多样性及遗传分化的影响 [D]. 上海：华东师范大学：6-7.]

HARTL DL, CLARK AG, 1989. Principles of population genetics [M]. 2nd ed. Sunderland：Sinauer Associates, Inc Publishers, 1990(1) : 128-129.

HUANG M, HUANG XP, LI BQ, et al., 2021. Optimization of *Liquidambar formosana* ISSR-PCR reaction systemand primer[J]. Mol Plant Breed, https://kns.cnki.net/kcms/detail/46.1068.s.20210429.1428.010.html. [黄敏，黄旭萍，李斌奇，等，2021. 枫香ISSR-PCR反应体系的优化及引物筛选[J]. 分子植物育种，https://kns.cnki.net/kcms/detail/46.1068.s.20210429.1428.010.html.]

LAI JX, 2020. Comparative analysis reveals different mechanisms of leaf color changes in *Liquidambar formosana* Hance. during different seasons and among different varieties [D]. Beijing: Chinese Academy of Forestry: 12-13. [赖玖鑫，2020. 枫香（*Liquidambar formosana* Hance.）叶色季节性变化及常彩叶机理研究 [D]. 北京：中国林业科学研究院：12-13.]

LI FF, 2015. Analysis on the genetic diversity of *Liquidambar formosana* of zheng jiang province [D]. Zhengzhou: Henan Agricultural University: 18-21. [李芳芳，2015. 浙江省枫香天然群体的遗传多样性研究 [D]. 郑州：河南农业大学：18-21.]

LIU FF, XU GB, LI TT, et al., 2021. Characterization and development of SSR markers from transcriptome sequence of *Bretschneidera sinensis* hemsl [J]. Plant Sci J, 39(3) : 288-296. [刘粉粉，徐刚标，李恬甜，等. 2021. 伯乐树转录组 SSR 分布特征及其引物开发 [J]. 植物科学学报，39(3) : 288-296.]

LIU JQ, FANG Y, CHEN YX, 2020. Analysis of SSR distribution characteristics and primer development of cornus florida transcriptome [J]. Mol Plant Breed, 18(20): 6769-6775. [刘佳奇，方彦，陈云霞，2020. 基于大花四照花转录组的SSR分布特征分析及引物开发 [J]. 分子植物育种，18(20): 6769-6775.]

LIU Z, ZHAO Y, YANG PD, et al., 2014. Comparison of parents identification for tea variety based on SSR, SRAP and ISSR markers [J]. J Tea Sci, 34(6):617-624. [刘振，赵洋，杨培迪，等，2014. SSR、SRAP、

ISSR分子标记在茶树品种亲本鉴定上的比较分析 [J]. 茶叶科学，34(6): 617-624.]
NEI M, 1973. Analysis of gene diversity in subdivided populations [J]. Proc Natl Acad Sci USA, 70(12) : 3321-3323.
XIN Y, CUI HR, ZHANG ML, et al., 2005. Development of EST(ex-pressed sequence tags)marker in Chinese cabbage and its transferability to rapeseed [J]. Hereditas, 27(3): 410-416. [忻雅，崔海瑞，张明龙，等，2005. 白菜的 EST 标记及其对油菜的通用性 [J]. 遗传，27(3): 410-416.]
SHANNON CE, WEAVE RW, WIENE RN, 1949. The mathematical theory of communication [J]. Urbana Univ Illinois Press, 85(2): 24-30.
SUN RX, 2017. Genetic diversity and phylogeography of *Liquidambar formosana* Hance in China [D]. Beijing: Chinese Academy of Forestry: 26-27. [孙荣喜，2017. 中国枫香树遗传多样性及谱系地理研究 [D]. 北京：中国林业科学研究院：26-27.]
WANG DD, YANG DX, 2017. Development and transferability of EST-SSR primers in *Actinidia arguta* [J]. J NW For Univ, 32 (4): 147-152. [王丹丹，杨东霞，2017. 软枣猕猴桃EST-SSR引物开发及通用性研究 [J]. 西北林学院学报，32(4): 147-152.]
WANG JJ， ZHANG B, ZHANG WS, 2015. A new variety，*Liquidambars formosana* 'Jinjue' [J]. Sci Sil Sin, 51(10): 154. [王建军，张波，张望舒，2015. 枫香新品种‘金钰’[J]. 林业科学，51(10): 154.]
WAN X, LIAO HH, HUANG HM, et al., 2013. Comparative test for DNA extraction methods of Citrus Huanglongbing pathogen [J]. J S Agric, 44(2):225-229. [王茜，廖惠红，黄宏明，等，2013. 柑橘黄龙病病原DNA提取方法比较 [J]. 南方农业学报，44(2); 225-229.]
WEI LD, LIANG LH, TANG SS, et al.，2016. Plantation Growth Process of *Liquidamba formosana* [J]. Guangxi For Sci, 45(4): 409-413. [韦理电，梁丽虹，唐生森，等，2016. 枫香人工幼林生长过程分析 [J]. 广西林业科学，45(4): 409-413.]
WEI LB, ZHANG HY, ZHENG YZ, et al.， 2008. Development and Utilization of EST-derivedmicrosatellites in Sesame (*Sesamum indicum* L.) [J]. Acta Agron Sin, 34(12):2077-2084. [魏利斌，张海洋，郑永战，等，2008. 芝麻EST-SSR标记的开发和初步研究 [J]. 作物学报，34(12)：2077-2084.]
XU ZJ, ZHAO S, XU L, et al., 2020. Discovery of microsatellite markers from RNA-seq data in cultivated peanut (*Arachis hypogaea*) [J]. China Agric Sci, 53(4):695-706. [徐志军，赵胜，徐磊，等，2020. 基于 RNA-seq 数据的栽培种花生 SSR 位点鉴定和标记开发 [J]. 中国农业科学，53(4):695-706.]
YANG GZ, ZHANG JB, ZHAO YM, et al., 2004. Study on polymorphism of 8 microsatellitemarkers in glassland red cattle crossed population[J]. Grass Feed Livestock, 124(3): 23-27. [杨国忠，张嘉保，赵玉民，等，2004. 8 种微卫星 DNA 标记在草原红牛杂交群体中多态性的研究[J]. 草食家畜，124(3)：23-27.]
YANG SP, SUN XM, WANG LH, et al., 2018. Development and identification of SSR markers based on transcriptome of *Helianthus tuberosus* L. [J]. Mol Plant Breed, 16(2): 484-492. [杨世鹏,孙雪梅,王丽慧,等,2018.基于菊芋转录组的 SSR 分子标记开发及鉴定 [J]. 分子植物育种，16(2): 484-492.]
YANG X, YANG N, YUAN QH, et al.，2021. Development and application of EST-SSR molecular markers in *Pinus bungeana* [J]. J Beijing For Univ, 43(7): 1-11. [杨雄，杨宁，袁启华，等，2021. 白皮松EST-SSR分子标记的开发及应用 [J]. 北京林业大学学报，43(7): 1-11]
YE P, LI XH, TANG JR, et al., 2019. Distribution and characteristics of SSR in transcriptome of *Camellia fascicularis* [J]. J Cent S Univ Technol, 39(9): 86-91. [叶鹏，李显煌，唐军荣，等，2019. 云南金花茶转录组 SSR 的分布及其序列特征 [J]. 中南林业科技大学学报，39(9): 86-91.]
YI LB, LI F, CUI ZX, 2017. Application of High-throughput sequencing technology in the research of ramie (*Boehmerianivea* L. Gaud) Biology [J]. Plant Fiber Sci Chin, 39(2): 88-93. [易浪波, 李富, 崔国贤, 2017. 高通量测序技术在苎麻生物学研究中的应用 [J]. 中国麻业科学，39(2): 88-93.]
YIN GP, LIU XS, JIANG Y, et al., 2021. The changes of tissue structure, photosynthetic characteristics and

pigment content of *Liquidambar formosana* leaves in the process of turning red [J/OL]. Guihai, https://kns.cnki.net/kcms/detail/45.1134.Q.20210722.1715.004.html. [尹国平，刘雄盛，蒋燚，等，2021. 枫香变红过程中叶片组织结构、光合特性及色素含量变化研究 [J/OL]. 广西植物，https://kns.cnki.net/kcms/detail/45.1134.Q.20210722.1715.004.html.]

ZHANG LD, TANG KX, 2010. Development of plant EST-SSR markers and its application [J]. Genom Appl Biol, 29 (3): 534-541. [张利达，唐克轩，2010. 植物EST-SSR标记开发及其应用 [J]. 基因组学与应用生物学，29(3): 534-541.]